

GenCore version 5.1.3  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 24, 2003, 10:19:03 : Search time 5267.53 Seconds  
(without alignments)  
11025.472 Million cell updates/sec

Title: US-09-922-895-3  
Perfect score: 3586  
Sequence: 1 GGATCCCTACCTTCCCATC.....CTATCAGAGGAGAGTGAA 3586

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues  
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*  
1: em\_estba:\*  
2: em\_esthum:\*  
3: em\_estin:\*  
4: em\_estmu:\*  
5: em\_estov:\*  
6: em\_estpl:\*  
7: em\_estro:\*  
8: em\_hic:\*  
9: gb\_est1:\*  
10: gb\_est2:\*  
11: gb\_hic:\*  
12: gb\_est3:\*  
13: gb\_est4:\*  
14: gb\_est5:\*  
15: em\_estfun:\*  
16: em\_estom:\*  
17: gb\_gss:\*  
18: em\_gss\_hum:\*  
19: em\_gss\_inv:\*  
20: em\_gss\_pln:\*  
21: em\_gss\_vrt:\*  
22: em\_gss\_fun:\*  
23: em\_gss\_mam:\*  
24: em\_gss\_mus:\*  
25: em\_gss\_Other:\*  
26: em\_gss\_pro:\*  
27: em\_gss\_trod:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
c 1	267.6	7.5	549	17	A0838273 HS_4712.A
c 2	210	5.9	453	10	B6677425 B6677425
c 3	210	5.9	700	17	AG114100 Pan trogl
c 4	209.6	5.8	450	17	A0035000 CIT-HSP-2
c 5	206.8	5.8	718	17	A0898001 HS_3135.A
c 6	203.8	5.7	352	10	AM089016 xd34d02.x

Result No.	Score	Query Match	Length	ID	Description
c 7	203.2	5.7	656	17	AG132086 Pan trogl
c 8	202.8	5.7	435	17	A0542219 RPT-11-3
c 9	202.2	5.6	660	17	AG157820 Pan trogl
c 10	202.2	5.6	739	17	A0035003 CIT-HSP-2
c 11	202	5.6	518	9	A1627168 A1627168
c 12	201.8	5.6	478	17	A0418548 RPT-11-2
c 13	201.8	5.6	782	12	BG743198 602634360
c 14	201.6	5.6	449	17	A0769842 HS_3170.B
c 15	201.4	5.6	628	14	BM716611 UT-E-DX1-
c 16	200.8	5.6	681	17	A0543621 RPT-11-3
c 17	200.6	5.6	808	12	BG432758 BQ432758
c 18	200.4	5.6	925	14	BQ433470 BQ433470
c 19	200.2	5.6	852	14	BQ690945 BQ690945
c 20	199.8	5.6	551	17	A0352725 CIT-D1-E1-
c 21	199.8	5.6	711	17	A0415030 RPT-11-2
c 22	199.6	5.6	521	17	A0237142 RPT-11-65
c 23	199.6	5.6	15970	17	A0839852 260L13-C5
c 24	199.4	5.6	518	17	A0412258 RPT-11-1
c 25	199.4	5.6	673	14	BM728941 UT-E-E01-
c 26	199.4	5.6	1024	14	BQ422616 BQ422616
c 27	199.2	5.6	478	17	BM7141 CIT-HSP-201
c 28	199.2	5.6	933	14	BQ959144 AGENCOURT
c 29	199	5.5	416	17	AQ140695 HS_3112-A
c 30	199	5.5	665	17	AG107923 Pan trogl
c 31	199	5.5	705	17	AG078610 Pan trogl
c 32	198.8	5.5	500	12	BM779049 BQ1464810
c 33	198.8	5.5	680	14	BM990843 UT-H-D10-
c 34	198.6	5.5	805	17	AQ739187 HS_5382-B
c 35	198.6	5.5	843	14	BQ421400 AGENCOURT
c 36	198.4	5.5	489	9	A1160786 A1160786
c 37	198.4	5.5	573	17	BA7427 CIT978SK-A-
c 38	198.4	5.5	674	17	AG123283 Pan trogl
c 39	198.2	5.5	466	14	BQ573431 UT-H-E21-
c 40	198.2	5.5	700	17	AG013777 Homo sapi
c 41	198.2	5.5	718	17	AG178110 Pan trogl
c 42	198.2	5.5	798	12	BG535971 602564142
c 43	198	5.5	879	14	BQ708582 AGENCOURT
c 44	198	5.5	2205	11	AF318346 Homo sapi
c 45	197.8	5.5	457	9	A1628859 A1628859

## ALIGNMENTS

RESULT 1  
A0838273/c 549 bp DNA linear GSS 30-AUG-1999  
DEFINITION HS\_4712\_A2\_H05\_SPEE CIT Approved Human Genomic Sperm Library D Homo sapiens genomic clone Plate-4712 COL-10 Row-O, DNA sequence.  
ACCESSION A0838273  
VERSION A0838273.1 GI:5808147  
KEYWORDS GSS.  
SOURCE human.  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 549)  
Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.  
TITLE Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome  
JOURNAL MEDLINE  
COMMENT Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)  
99380589  
Contact: Mahairas GG, Wallace JC, Hood L  
High Throughput Sequencing Center  
University of Washington  
401 Oudeen Anne Avenue North, Seattle, WA 98109, USA  
Tel: (206) 616-3618  
Fax: (206) 616-3887  
Email: jwallace@u.washington.edu  
Clones may be purchased from Research Genetics (info@resgen.com).  
BAC end Web Server: http://www.htsc.washington.edu



JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 700)
AUTHORS	Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T., Tokoki,Y., Watanabe,H. and Sakaki,Y.
TITLE	Direct Submission
JOURNAL	Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Suehiro-cho,Tsurumi-ku, Yokohama, Kanagawa 220-0045, Japan (E-mail:chimbpgsc@riken.go.jp, URL:http://npg.gsc.riken.go.jp/, Tel.:81-45-503-9111, Fax:81-45-503-9170) Clones are derived from the chimpanzee BAC library pYB This BAC end was generated during the R&D process and may have higher chance of clone tracking errors.
COMMENT	PRIMERS  Sequencing: M13Rev  LIBRARY  Vector : pRS145 R.Site 1 : SacI R.Site 2 : SacI  Location/Qualifiers  1..700 /organism="Pan troglodytes" /db_xref="taxon:9598" /clone="PMB-120022.R" /sex="male" /cell_type="lymphoblast" /clone_id="PYB Chimpanzee Male BAC Library"
BASE COUNT	200 a 167 c 176 g 157 t
ORIGIN	
Query Match	5.9% Score 210; DB 17; Length 700;
Best Local Similarity	81.9%; Pred.No.1,5e-22;
Matches 254; Conservative	0; Mismatches 55; Indels 1; Gaps 1;
OY 1079	TTATTTATTATTAATTTATTTTATTTTGAGATGAGCTGGCTGCGCCAGCGGTGAGTGC 1138
Db 494	TTATTTATTATTTATTTA-TTTTGAACAACATCTCTCTGTCCGCCGGCTGAGATGC 436
OY 1139	AGCGCGGTGATCACAGTCACACTGCAGCCTCAACCTTAGCGTCAAGGAAATTCGCCACC 1198
Db 435	TGTGATGTGATCTCGCTCAGTCAGCCTCTCATCTCTGAGCTCAAGTATCTCCACC 376
OY 1199	TCAGCCCCCAAGTAGTGGGACCAACAGTATGGCCCACATGGCTGGCTAATTTCTTAT 1258
Db 375	TCAGCCTCCCAAGTAGTGGGACACAGGACAGTCGCCACATGCCCTGGATGATTTGTGA 316
OY 1259	TTTTTTAGAGAATAGGATGCACATAATGTGCGACGGCTGGCTTGAATCCTGGGCTCA 1318
Db 315	ACTTTTGTAGAGACACAGTTTACATATTTGTGCCAGGTTGGTCTTGAACCTCTGGCCTCA 256
OY 1319	GGTGAGCCTCCACACCTGGGCTCCCAAAGTACTGAGGATTAACAGGCATGAGCCAGTCCC 1378
Db 255	AGTGAATCCTCCCTCAGCCTCCCAAAGGTGGATTAACAGGGGTGACACAGCACACC 196
OY 1379	CTGCCCATAT 1388
Db 195	CAGCCAGGAT 186
RESULT 4	
A0035000	
LOCUS	A0035000 450 bp DNA linear GSS 11-JUL-1998
DEFINITION	CIT-HSP-2333P5."TF CIT-HSP Homo sapiens genomic clone 2333P5, DNA sequence.
ACCESSION	A0035000
VERSION	A0035000.1 GI:3301174
KEYWORDS	GSS.
SOURCE	human.
ORGANISM	Homo sapiens
Eukaryote:	Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE	1 (bases 1 to 450) Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,

[illegible]

	ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrates; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
	REFERENCE	1 (bases 1 to 718)
	AUTHORS	Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and Hood,L.
TITLE	Sequence-tagged connectors: A sequence approach to mapping and scanning the human genome	
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)	
MEDLINE	99380589	
COMMENT	Contact: Mahairas GG, Wallace JC, Hood L High Throughput Sequencing Center University of Washington 401 Queen Anne Avenue North, Seattle, WA 98109, USA Tel.: (206) 616-3618 Fax: (206) 616-3887 Email: jwallace@u.washington.edu Clones may be purchased from Research Genetics (info@resgen.com). BAC end Web Server: http://www.hisc.washington.edu Plate: 3135 row: O column: 14 Seq primer: f7 Class: BAC ends High quality sequence stop: 718. Location/Qualifiers  1..718 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="plate=3135 col=14 Row=O" /clone.lib="CIT Approved Human Genomic Sperm Library D" /sex="male" /note="Organ: sperm; Vector: pBelBAC11; BAC Clones in E-Coli DH10B"	
BASE COUNT	149 a     157 c     122 g     287 t       3 others	
ORIGIN		
Query Match	5.8%; Score 206.8; DB 17; Length 718;	
Best Local Similarity	74.0%; Pred. No. 4.5e-22;	
Matches 262; Conservative	0; Mismatches 92; Indels 0; Gaps 0;	
Oy	1056 GTTTTAAATTTTTTAATTATAATTATTTATTTATTTATTTTGAGTAGACTCT 1115 	
Dd	194 GTTTAAGATTTATTTATTTATTTATTTATTTATTTATTTTCGACAGAATCC 253	
Oy	1116 GCCTGTGCCAGCGGTGGAGTGCACGGCGGTGATCACAGTTCACGTCAACCTTC 1175 	
Dd	254 CACCTCTTCCCAGCCAGCATAGTACAGTGGCGGCATCGCGCTCACGAACAATTGGCTCC 313	
Oy	1176 TAGGCTCAAGGATTTCTCCACCACAGCCCCCAAAGTAGTTGGACACAGCATATGCCC 1235 	
Dd	314 TGGATTCAAGCCCTTCTCGCTCAGCCCTCCCGAGTACTGTATTACAGCGGTGCACC 373	
Oy	1236 ACCATGCGCTGGGTAAATTTCTTATTTTGTGTAGACATGAGTCACTATATTGTCCAAG 1295 	
Dd	374 ACCTATGCGCTGGGTAAATTTTGTATTTTGTGTATTTTGTGTAGACATGAGTCACTATATTGTCCAAG 433	
Oy	1296 CTGGTCTTGAATTTCTGGGCTCAGSTGAGCCCTCCACACCTGGGCTCCCAAGATAGTGGA 1355 	
Dd	434 CTGGCTCAAACTCGTACCTCAAGTAGATCTCCGCGCTGGCCTGCCAACAAATGCTGGGA 493	
Oy	1356 TTACAGGCGATGAGCCAAAGTCCCTCGCCCATATGTAGATTTTGTGTCTGTATCC 1409 	
Dd	494 TTACAGGCTGTAGACATCATGCTCGGCGCTGTATTATGATTTTCTTTATGATGCC 547	
RESULT 6		
AW089016/c		
LOCUS	AM089016                  352 bp         mRNA         linear         EST 15-OCT-1999	
DEFINITION	xcd34d02.x1 NCL_GCAP.Ov23 Homo sapiens cDNA clone IMAGE:2595651 3'	
	similar to contains Alu repetitive element::contains element HGR	
	repetitive element :: mRNA sequence.	
ACCESSION	AW089016	
VERSION	AW089016.1 GI:6044851	

KEYWORDS	EST.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. 1 (bases 1 to 352)
AUTHORS	NCI-CCAP <a href="http://www.ncbi.nlm.nih.gov/ncicgap">http://www.ncbi.nlm.nih.gov/ncicgap</a> .
TITLE	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index Unpublished (1997)
JOURNAL	Contact: Robert Strausberg, Ph.D. Email: cgaps@email.nih.gov Tissue Procurement: Christopher Moskalko, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
COMMENT	CDNA Library Preparation: Life Technologies, Inc. DNA Library Arrayed by: Greg Lennon, Ph.D. DNA Sequencing by: Washington University Genome Sequencing Center clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <a href="http://www.bio.lnl.gov/bdrip/Image/Image.html">www.bio.lnl.gov/bdrip/Image/Image.html</a> Seq primer: -400P from Gdbco High quality sequence stop: 321.
FEATURES	Location/Qualifiers
SOURCE	1..352 /organism="Homo sapiens" /db_xref="taxon:9606" /cdna_xref="IMAGE:2595651" /clone_lib="NCI CGAP_OV23" /tissue_type="tumor, 5 pooled (see description)" /lab_host="DH10B" /note="Organ: ovary; Vector: PCW-SPOrt6, Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT. Average insert size 1.35 kb. Tumor types include: mixed Mullerian tumor, papillary serous, clear cell, spindle cell. All are primary tumors, metastasis positive. Life Technologies catalog #: 11534-013"
BASE COUNT	115 a       81 c       82 g       74 t
ORIGIN	
Query Match	5.7%; Score 203.8; DB 10; Length 352;
Best Local Similarity	76.5%; Pred. No. 1.8e-21;
Matches 250:	Conservative 0; Mismatches 77; Indels 0; Gaps 0;
Db	1072 TTTTAAATTAATTATTTATTTATTTATTTAGAGATGGAGCTGGCTGCCAGGCT 1131       Db 337 TATGATATGATGTATGATTTATTTATTTAGAGAGATCTTGCTGTCAACCAGGCT 278       QY 1132 GGAGTCAGCGCGGTGATACACAGTTCACTGCAGGCTCAACCTTCTAGGCTCAAGGATTC 1191       Db 277 GGAGTCAGTCGGCACGATCTGGCGCTACACACTCTGCTCTGGGTTCAGATTC 218       QY 1192 TCCCACCTCAGCCCCCAAGTAGTTGGACACACGATATGGCCACCATGGCTGGCTAAT 1251       Db 217 TCCCACCTCAGCTCTCTAGTACCTGGGATTCACAGCATGGCCACGACTCTGGCTAAT 158       QY 1252 TTCTTAATTTTTTTGTAGAGATAGATCTCACTATATTGTCCAGGCTGGTCTTGAAATTCCT 1311       Db 157 TTTTGTATTTTATAGAGATGGGGTTTCCACATGTTTACCAAGGTGTTTGATAACT 98       QY 1312 GGGCTCAGGTAGGCTCCGACCTGGGCTCCCAAAGTACTGGGATTACAGCATGAGCCA 1371       Db 97 GACCTCAGGTAGTCCACCCACCTGGCTCCCAAAGTCTGGGATTACAGCATGAGCCA 38       QY 1372 AGGTCCCTGGCCATATGAGATTTTCT 1398       Db 37 CTGTGCCCGGCTCATTTTCTTTTCTT 11
RESULT 7	
LOCUS	AG132086 656 bp DNA linear GSS 04-NOV-2001
DEFINITION	Pan troglodytes DNA, clone: PTB-144E15.R, genomic survey sequence.
ACCESSION	AG132086

[illegible]

DEFINITION	RPC1-11-36J1J3.TV RPC1-11 Homo sapiens genomic clone RPC1-11-36J1J3 , DNA sequence.
ACCESSION	AQ542219
VERSION	AQ542219.1 GI:4872675
KEYWORDS	GSS.
SOURCE	human.
ORGANISM	Homo sapiens
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo. 1 (bases 1 to 435)
AUTHORS	Zhao,S., Adams,M.D., Nierman,W., Malek,J., de Jong,P. and Venter, J.C.
TITLE	Use of BAC End Sequences from Library RPC1-11 for Sequence-Ready Map Building
JOURNAL	Unpublished (1997) Other GSSs: RPC1-11-36J1J3.TV
COMMENT	Contact: Shaying Zhao, William Nierman, Mark Adams Department of Eukaryotic Genomics The Institute for Genomic Research 9712 Medical Center Dr., Rockville, MD 20850 Tel.: 301 838 0200 Fax: 301 838 0208 Email: hbe@tigr.org Clones are derived from the human BAC library RPC1-11. For BAC library availability, please contact Pieter de Jong (pietere@jeng.med.buffalo.edu). Clones may be purchased from BACPAC Resources ( <a href="http://bacpac.med.buffalo.edu/ordering">http://bacpac.med.buffalo.edu/ordering</a> ) or from Research Genetics cs ( <a href="mailto:info@resgen.com">info@resgen.com</a> ). BAC end search page: <a href="http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html">http://www.tigr.org/tdb/humgen/bac_end_search/bac_end_search.html</a> . Seq primer: SP6 Class: BAC ends.
FEATURES	Location/Qualifiers 1..435 /organism="Homo sapiens" /db_xref="GDB:7638468" /db_xref="taxon:9606" /clone="RPC1-11-36J1J3" /clone_1id="RPC1-11" /sex="Male" /cell_type="Lymphocytes" /note="Vector: pBAC3.6; Site_1: EcoRI; Site_2: EcoRI; RPC111 Human Male BAC Library"
BASE COUNT	87 a     110 c     101 g     137 t
ORIGIN	
Query Match	5.7%; Score 202.8; DB 17; Length 435;
Best Local Similarity	74.9%; Pred. No.2.3e-21;
Matches 268:	Conservative 0; Mismatches 87; Indels 3; Gaps 1;
OY 1039	GGTGAAGAACAATCATAACGGTTTAAATTTTTTAATTAAATTTAATTTAATTTAAT 1098
Dd 41	GGTAGACTCTCCACGCTTTTATTATTATTATTATTATTATTATTATTATTATTATT 97
OY 1099	TTTTTGAGATGSAGTCCTGGCTGCCTGCCCAAGCGTGGAGTGCACGGCGTATCACAGTTCA 1158
Dd 98	TTTGGACACAGAGTCCTGCTGTGTGCCCCAGCGCTGGCGTAGTCAGATCATCTCAGCTCA 157
OY 1159	CTGCAGCGCTCAACCTTTCAGAGCTCAAGGAGTTTCGCCACCTCAGCCCCCAAGTAGTGG 1218
Dd 158	CTGCACACTCCACCCTCCCGGGGTCAAGCAAATTTCTCTGCTCAGCTCCCAAGTAGTGG 217
OY 1219	GACCAACGTAATGGCGCCACCAATGCTCGCTAAATTTCTAATTTTGTGAGATAGATTC 1278
Dd 218	GACCAACAGTGTGGCCACCAACATGCTGCTAATTTTTCATATTTTATAGTAGAGACGGGTTT 277
OY 1279	TCACTAATATGTCCAGCGTGTCTTGAATTCCTGGCGTCAAGTGAAGCTCCACCTGGGC 1338
Dd 278	TCACCATTTTGGCGAGCGTGTTCACCAATCTCCTGACCTCAGAGTGATTCACCCACCTTGGC 337
OY 1339	CTCCCAGTAATCGGAGTTACAGCATATGACCAAGTCCCTGCCCATATAGATTTT 1396
Dd 338	CTCCCAGTCTGGAGATTACAGGTGATGATGACGCCCCACACTTTCATCGAGCTTTT 395

```
RESULT 9
AG157820
LOCUS
DEFINITION
Pan troglodytes DNA, clone: RP43-022N18.TJ, genomic survey
sequence.
ACCESSION
AG157820
VERSION
AG157820.1 GI:16687498
KEYWORDS
GSS.
SOURCE
Pan troglodytes male lymphocytes DNA, clone:lib:RPCT-43 Chimpanzee
Male BAC Library clone:RP43-022N18.TJ.
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Pan.
REFERENCE
1
Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
Totohi,Y., Watanabe,H. and Sakaki,Y.
BAC end sequences of library RPCT-43
JOURNAL
Unpublished
AUTHORS
2 (bases 1 to 660)
Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
Totohi,Y., Watanabe,H. and Sakaki,Y.
TITLE
Direct Submission
JOURNAL
Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
1-7-22 Suehiro-Chou,Tsuriuni-Ku, Yokohama, Kanagawa 230-0045, Japan
(E-mail:chimbse@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
Tel:81-45-503-9111, Fax:81-45-503-9170)
COMMENT
Clones are derived from the chimpanzee BAC library RPCT-43 This BAC
end was generated during the R&D process and may have higher chance
of clone tracking errors.
PRIMERS
Sequencing: TJ
LIBRARY
Vector : PACe3.6
R.Site 1 : EcoRI
R.Site 2 : EcoRI
FEATURES
source
1..660
Location/Qualifiers
/organism="Pan troglodytes"
/db_xref="taxon:9598"
/clone="RP43-022N18.TJ"
/sex="male"
/cell_type="lymphocytes"
/contig_id="RPCT-43 Chimpanzee Male BAC Library"
BASE COUNT
166 a 154 c 135 g 205 t
ORIGIN
Query Match 5.6%; Score 202.2; DB 17; Length 660;
Best Local Similarity 74.3%; Pred. No. 2.3e-21;
Matches 255; Conservative 0; Mismatches 88; Indels 0; Gaps 0;
```

```
Db 359 TTATAGCATGAGCAGCAGTGTCCCGCCCATTTATTTGTTAT 401
|||||
RESULT 10
AQ035003/c
LOCUS
DEFINITION
A0035003
CIT-HSP-2333P7.TR CIT-HSP Homo sapiens genomic clone 2333P7, DNA
sequence.
ACCESSION
AQ035003
VERSION
AQ035003.1 GI:3301177
KEYWORDS
GSS.
SOURCE
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 739)
Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
Simon,M. and Venter,J.C.
Use of a random BAC End Sequence Database for Sequence-Ready Map
Building (1998)
JOURNAL
Unpublished (1998)
COMMENT
Other_GSSs: CIT-HSP-2333P7.TP
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@ligr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.ligr.org/tldb/hungen/bac_end_search/bac_end_search.html.
Seq primer: M13 Reverse
Class: BAC ends.
FEATURES
source
1..739
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="2333P7"
/contig_id="CIT-HSP"
/sex="Male"
/cell_type="Sperm"
/contig_id="pbelobac11; site_1: HindIII; site_2:
HindIII"
BASE COUNT
246 a 147 c 175 g 171 t
ORIGIN
Query Match 5.6%; Score 202.2; DB 17; Length 739;
Best Local Similarity 74.3%; Pred. No. 2.2e-21;
Matches 255; Conservative 0; Mismatches 88; Indels 0; Gaps 0;
```

OY	1359	CAGGCATGACCAAGTCCCGGCCCATATGAGATTTCCTGNC	1401
Dd	255	CAGGCATGAGCCACTGACACTAAGCCATTTCACCACCTCTATC	213
RESULT	11		
LOCUS	A1627168	478 bp	mRNA linear EST 07-MAR-2000
DEFINITION	ly06a05.x1 NCI-CGAP utr2 Homo sapiens cDNA clone IMAGE:2283440 3'		
ACCESSION	A1627168		
VERSION	A1627168.1	GI:4663968	
KEYWORDS	Esf.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	1 (bases 1 to 478)		
JOURNAL	NCI-CGAP http://www.nccl.nlm.nih.gov/ncicgap.		
COMMENT	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index unpublished (1997)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgapbs-remail.nih.gov		
	Tissue Procurement: Christopher Moskalko, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.		
	cDNA Library Preparation: Life Technologies, Inc.		
	cDNA Library Arrayed by: Greg Lennon, Ph.D.		
	DNA Sequencing by: Washington University Genome Sequencing Center		
	clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LNLW at:		
	www.bio.lnl.gov/bdrp/image/image.html		
	Insert Length: 207 Std Error: 0.00		
	Seq primer: -40UP from Gibco		
	High quality sequence stop: 417		
FEATURES	POLYA-No.		
SOURCE	Location/Qualifiers		
	1..478		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:2283440"		
	/clone_lib="NCI-CGAP-utr2"		
	/cuisine_type="moderately-differentiated endometrial adenocarcinoma, 3 pooled tumors"		
	/lab_host="DH10B"		
	/note="Organ: uterus; Vector: pCMV-SPORT6; Site:1: SalI; Site:2: NotI; Cloned unidirectionally. Primer: oligo dT. Average insert size 1.85 kb. Life Technologies catalog #: 11539-012"		
BASE COUNT	118 a 118 c 95 g 147 t		
ORIGIN			
Query Match	5.6%; Score 202; DB 9; Length 478;		
Best Local Similarity	72.4%; Pred. No. 2.9e-21;		
Matches 262; Conservative	0; Mismatches 100; Indels 0; Gaps 0;		
OY	1075	TTAATTATTATTATTATTATTATTATTATTGAATGGAGTGCGTGGCCCCAGGCTGGA	1134
Dd	1	TTAATTATTATTATTATTGTGAGACAGCTTTTTTTGAAATCTCACTCTGTGCCCCAGGCTGGA	60
OY	1135	GTCGACGGCGGTGATCACAGTTCACCTGCAGCCTCAACCTTAGGCTCAAGGATTCTCC	1194
Dd	61	GTCGATGGCGCTGATCTGGCTGACCTGCACCACTGAGCHCTCGGTTCAAGGATTCTCC	120
OY	1195	CACGTCAGGCCCCCAAGTAGTTGGGACCAACAGTAGTGCGCCACATGCTGGCTAATTTC	1254
Dd	121	TGCGTCAGGCTCCCAAGTAGCTGGGACACTACAAGCATGCCCATGCTGCTAATGTT	180
OY	1255	TTAATTATTATTAGAAGATGAGATCTACATATTTTCCAGGCTGGCTGAATTCCTGGG	1314
Dd	181	TGTATTATTTTAGTAGACAGAGGTTTTCGCATATTGGCCACAGGCTGGTCTCGAATCTCTGGC	240
OY	1315	CTCAGGTGAGCCCTCCACCTTGCGGCTCCCAAAGTAAGTCTGGAGTTACAGGCATGAGCCAAG	1374

[illegible]

```

Db 129 ACCTGCGCTCCCGGGGTCAAGAGATTCTTGCCCTCAGCCTCTGAGTAGCTGGAGCTA 188
Oy 1224 CACGATGGCCACCATGCTGGCTTAATTTCTATTTTGTAGAGATAGATCTCACT 1283
Db 189 CAGGCGCGTGCAACCGCTGGCTGCTAATTAATTAATTAATTAATTAATTAATTAAT 248
Oy 1284 AATATGTCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1343
Db 249 ATGTTGGCAGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 308
Oy 1344 AAAGTACTGGGATTAACAGCAGTCCAGGCTCCCTGCC 1384
Db 309 AAAGTCTGGGATTAACAGCAGTATTCACCGAGCCAGCC 349

```

```

RESULT 13
Bg743198      782 bp      mRNA      linear      EST 15-MAY-2001
LOCUS      Bg743198/c
DEFINITION      602634360F1 NCI_CGAP_Skn3 Homo sapiens cDNA clone IMAGE:479326 5',
                mRNA sequence.
ACCESSION      Bg743198
VERSION      Bg743198.1 GI:14053851
KEYWORDS      EST.
SOURCE      human.
ORGANISM      Homo sapiens
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE      1 (bases 1 to 782)
                NIH-MGC http://mgi.nci.nih.gov/.
                National Institutes of Health, Mammalian Gene Collection (MGC)
                JOURNAL      Unpublished (1999)
                COMMENT      Contact: Robert Strausberg, Ph.D.
                Email: cgaab-r@mail.nih.gov
                Tissue Procurement: James Cleaver, M.D.
                cDNA Library Preparation: Life Technologies, Inc.
                Sequencing by: Incyte Genomics, Inc.
                Clone distribution: MGC clone distribution information can be
                found through the I.M.A.G.E. Consortium/LLNL at:
                http://image.llnl.gov
                Plate: LLNL0636 row: h column: 15
                High quality sequence stop: 520.
                Location/Qualifiers
                1. 782
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone_id="IMAGE:479326"
                /lab_host="NCI_CGAP_Skn3"
                /note="Organ: skin; Vector: PCMV-SPORT6; Site: 1: NCI;
                Site: 2: Sali; Cloned unidirectionally. Primer: Oligo dt.
                Average insert size 1.5kb. Library constructed by Life
                Technologies. Note: this is a NCI_CGAP Library."

```

```

FEATURES
    source
        1. 782
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone_id="IMAGE:479326"
        /lab_host="NCI_CGAP_Skn3"
        /note="Organ: skin; Vector: PCMV-SPORT6; Site: 1: NCI;
        Site: 2: Sali; Cloned unidirectionally. Primer: Oligo dt.
        Average insert size 1.5kb. Library constructed by Life
        Technologies. Note: this is a NCI_CGAP Library."

```

```

BASE COUNT      230 a      196 c      184 g      172 t
ORIGIN
Query Match      5.6%; Score 201.8; DB 12; Length 782;
Best Local Similarity 72.1%; Pred. No. 2.5e-21;
Matches 263; Conservative 0; Mismatches 102; Indels 0; Gaps 0;

```

```

Oy 1041 TGAAGACATATCAGGTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 1100
Db 542 TGAATTTGTTCCCGGGTGGGTTTATTTTATTTTATTTTATTTTATTTTATTTT 483
Oy 1101 TTGAGATGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1160
Db 482 TTGAGACACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 423
Oy 1161 GCAGCCCAACCTTCTGAGTCAAGGATTCCTCCACCTGAGCCGCCCAAGTGTGGGA 1220
Db 422 GCAGCCTCTGCTGCCAGGTTCAAGCGATTCCTGCTGAGCCTCCCGAGTGGGA 363

```

```

Oy 1221 CCACAGTATGGCCACCATGCTGGCTTAATTTCTATTTTGTAGAGATAGATCTC 1280
Db 362 TTACAGGACGATGCCACCATGCTGGCTTAATTTCTATTTTGTAGAGATAGGAGTTT 303
Oy 1281 ACTATTTTGTCCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1340
Db 302 ACCATGTTGGCCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 243
Oy 1341 CCCAAGTACTGGGATTAACAGCAGTCCAGGCTCCCTGCCATATGAGATTTTCTGT 1400
Db 242 CCCAAGTCTGGGATTAACAGCAGTCCAGGCTCCCTGCCATATGAGATTTTCTAC 183
Oy 1401 CTCGTG 1405
Db 182 ATCTG 178

```

```

RESULT 14
A0769842      449 bp      DNA      linear      GSS 28-JUL-1999
LOCUS      A0769842
DEFINITION      HS.3170.B2.G07.T7C CIT Approved Human Genomic Sperm Library D Homo
                sapiens genomic clone Plate=3170 Col=14 Row=N, DNA sequence.
ACCESSION      A0769842
VERSION      A0769842.1 GI:5647958
KEYWORDS      GSS.
SOURCE      human.
ORGANISM      Homo sapiens
                Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
                Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE      1 (bases 1 to 449)
                Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holman,T.,
                Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
                Hood,L.
                Sequence-tagged connectors: A sequence approach to mapping and
                scanning the human genome
                Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
                MEDLINE      99380589
                JOURNAL      Contact: Mahairas GG, Wallace JC, Hood L
                High Throughput Sequencing Center
                University of Washington
                401 Queen Anne Avenue North, Seattle, WA 98109, USA
                Tel: (206) 616-3618
                Fax: (206) 616-3867
                Email: jwallace@u.washington.edu
                BAC end Web Server: http://www.htsc.washington.edu
                Plate: 3170 row: N column: 14
                Seg primer: T7
                Class: BAC ends
                High quality sequence stop: 449.
                Location/Qualifiers
                1. 449
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone_id="Plate=3170 Col=14 Row=N"
                /clone_lib="CIT Approved Human Genomic Sperm Library D"
                /sex="male"
                /note="Organ: sperm; Vector: pBelobAC11; BAC Clones in
                E-Coli DH10B"

```

```

FEATURES
    source
        1. 449
        /organism="Homo sapiens"
        /db_xref="taxon:9606"
        /clone_id="Plate=3170 Col=14 Row=N"
        /clone_lib="CIT Approved Human Genomic Sperm Library D"
        /sex="male"
        /note="Organ: sperm; Vector: pBelobAC11; BAC Clones in
        E-Coli DH10B"

```

```

BASE COUNT      78 a      114 c      96 g      158 t      3 others
ORIGIN
Query Match      5.6%; Score 201.6; DB 17; Length 449;
Best Local Similarity 75.4%; Pred. No. 3.4e-21;
Matches 264; Conservative 0; Mismatches 84; Indels 2; Gaps 1;

```

```

Oy 1055 GGTATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTTAAATTTT 1114
Db 100 GCTTGTATATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 159
Oy 1115 TGG--CTGTGCCAGGCTGAGTGCAGCGCGGTGATCAAGTTCACTGAGGCTCAAC 1172
Db 160 TTGCTGTGTTGCCAGCTGAGTGCAGCGCGGTGATCAAGTTCACTGAGGCTGCTGCC 219

```





